The Need and Requirements for Validating Damage Detection Capability

E. A. LINDGREN and C. F. BUYNAK

ABSTRACT

Structural Health Monitoring (SHM) has been proposed to benefit the maintenance cost and reliability of aircraft structures through the early detection of damage during service and in support of condition-based maintenance. At the same time, if the integrity of an aircraft component is dependent upon the performance of an SHM system to detect damage, which has recently been referred to as an *in-situ* NDE system by two airframe OEMs [1,2], the reliability of the validated capability of the SHM system must also be ensured over the service life of the aircraft. To enable the calculation of risk, which is the metric used by the US Air Force (USAF) to manage the integrity of structures, the capability of any inspection process must be assessed and integrated into the risk calculation. Thus, a qualification plan is required for both the validated capability, i.e. POD and false-positive rate, and the durability of the sensing technique. These requirements were established by the Aircraft Structural Integrity Program (ASIP) Senior Leader for the USAF, Mr. Charles Babish [3]. Empirical assessment of the performance of an SHM system in such an environment is not a trivial manner and could readily be cost prohibitive. A protocol has been developed at the request of the AFRL Team Leader for Integrated System Health Assessment initiative that leverages current research and development efforts to minimize the amount of empirical data required for assessing the Probability of Detection (POD) of a damage detection system [4]. The on-going effort uses models and transfer functions to minimize the need for empirical data and is called Model-assisted Probability of Detection, or MAPOD [5,6]. The details defining these requirements and structure for the protocol to satisfy these requirements are presented.

INTRODUCTION

With the emergence of techniques to permanently mount damage detection sensors on structures, which was first used as a method to detect cracks on USAF

Eric A. Lindgren, Charles F. Buynak, Materials and Manufacturing Directorate, Air Force Research Laboratory, WPAFB, Ohio 45433, USA;

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE **SEP 2011**	2. REPORT TYPE **N/A**	3. DATES COVERED **-**
4. TITLE AND SUBTITLE **The Need and Requirements for Validating Damage Detection Capability**		5a. CONTRACT NUMBER
		5b. GRANT NUMBER
		5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)		5d. PROJECT NUMBER
		5e. TASK NUMBER
		5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Materials and Manufacturing Directorate, Air Force Research Laboratory, WPAFB, Ohio 45433, USA;**		8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)		10. SPONSOR/MONITOR'S ACRONYM(S)
		11. SPONSOR/MONITOR'S REPORT NUMBER(S)
12. DISTRIBUTION/AVAILABILITY STATEMENT **Approved for public release, distribution unlimited**		
13. SUPPLEMENTARY NOTES **See also ADA580921. International Workshop on Structural Health Monitoring: From Condition-based Maintenance to Autonomous Structures. Held in Stanford, California on September 13-15, 2011 . U.S. Government or Federal Purpose Rights License.**		

14. ABSTRACT

Structural Health Monitoring (SHM) has been proposed to benefit the maintenance cost and reliability of aircraft structures through the early detection of damage during service and in support of condition-based maintenance. At the same time, if the integrity of an aircraft component is dependent upon the performance of an SHM system to detect damage, which has recently been referred to as an in-situ NDE system by two airframe OEMs [1,2], the reliability of the validated capability of the SHM system must also be ensured over the service life of the aircraft. To enable the calculation of risk, which is the metric used by the US Air Force (USAF) to manage the integrity of structures, the capability of any inspection process must be assessed and integrated into the risk calculation. Thus, a qualification plan is required for both the validated capability, i.e. POD and false-positive rate, and the durability of the sensing technique. These requirements were established by the Aircraft Structural Integrity Program (ASIP) Senior Leader for the USAF, Mr. Charles Babish [3]. Empirical assessment of the performance of an SHM system in such an environment is not a trivial manner and could readily be cost prohibitive. A protocol has been developed at the request of the AFRL Team Leader for Integrated System Health Assessment initiative that leverages current research and development efforts to minimize the amount of empirical data required for assessing the Probability of Detection (POD) of a damage detection system [4]. The on-going effort uses models and transfer functions to minimize the need for empirical data and is called Model-assisted Probability of Detection, or MAPOD [5,6]. The details defining these requirements and structure for the protocol to satisfy these requirements are presented.

15. SUBJECT TERMS

<table>
<tr><td colspan="3">16. SECURITY CLASSIFICATION OF:</td><td rowspan="2">17. LIMITATION OF ABSTRACT
SAR</td><td rowspan="2">18. NUMBER OF PAGES
8</td><td rowspan="2">19a. NAME OF RESPONSIBLE PERSON</td></tr>
<tr><td>a. REPORT
unclassified</td><td>b. ABSTRACT
unclassified</td><td>c. THIS PAGE
unclassified</td></tr>
</table>

aircraft in the mid and late 1980s [7], there has not been a formal process to assess the capability to detect damage. However, if this class of permanently mounted damage detection techniques is to be used on USAF aircraft for detection of damage in structural components, this capability must be validated to enable the output from the sensing method to be integrated into the ASIP, which is the method the USAF uses to maintain the integrity of aircraft structure. Current ASIP methods, as defined in MIL-STD-1530C, use a damage tolerance approach after the F-111 wing separation that occurred in 1968 after approximately 100 hours of flight for an aircraft that was qualified for 4000 hours. The metric by which integrity of the aircraft structure is measured is risk of structural failure per flight hour and the threshold of the acceptable risk is $1x10^{-6}$ with an objective of $1x10^{-7}$. Overviews of the current fleet indicate that the risk is lower than the threshold of acceptable risk, which points to the success of the current ASIP process [3].

A key component of the damage tolerance approach is periodic inspection. To enable the calculation of risk, the capability of the inspection process must be validated and integrated into the risk calculation. The method to assess the capability of a damage detection technique is known as probability of detection (POD) as the entire POD curve is used in the risk calculation. The guidance to determine POD is given by MIL-HDBK-1823A [6]. Revision A of this handbook is a recent update that includes the potential to use model-assisted methods to determine POD. Note that the risk calculations used by ASIP require a probability of detection curve and not a point estimate of capability.

Therefore, an in-situ damage detection method applied to any ASIP managed structural feature, commonly referred to as fatigue critical locations and/or control points, must have a POD curve for the relevant structure that is being inspected to enable the capability to be integrated into the current approach for force structural maintenance plans (FSMP). To seek alternative methods to validate the capability of an in-situ damage detection method would require a change in how the integrity of the structure is maintained, which would represent even higher barriers of acceptance when compared to the challenges of generating a POD curve for an in-situ damage detection method

Similar requirements emerge when considering future maintenance process, such as Condition-based Maintenance (CBM), which require the assessment of the condition of the structure. For these applications, the capability must be defined in terms of location in x, y, and z, as well as the three dimensions of the damage or region of interest.

In addition to the capability of the system, the rate of false positives and the durability of the sensing system must be quantified as a function of time and usage [6]. All these parameters are required to enable a cost benefit analysis to be performed to determine if the in-situ damage sensing system could be used on USAF aircraft. Thus, a qualification plan is required for both the capability, i.e. POD and false-positive rate, and the durability of the sensing technique. The details defining these requirements and structure for the protocol to satisfy these requirements are presented and will include several examples of possible applications and representative examples when these requirements are not quantitatively satisfied.

DRIVERS FOR VALIDATION PROTOCOL

Thus, if the output from an in-situ damage detection system is to be used in the framework of the USAF ASIP, the capability of an SHM system to perform its function must be validated over the range of expected damage states including the likely extreme loading and environmental spectrum to which the system will be exposed. For in-situ damage detection methods aimed at damage detection, the ASIP Senior Leader has stated that a POD curve for the detection of damage is required [3] to establish that a technique can reliably detect damage of a critical size. It is very possible that other services and the commercial airline community could have different requirements for the validation process as these entities will use alternative methods to ensure structural integrity. Thus, the guidelines for the methods to validate the capability of an in-situ damage detection method could be different from the USAF interests. In addition, if the structural region of interest being assessed does not fall into the category of being managed according to ASIP, then there could be alternative options, but the requirements for validation are still established by the ASIP community.

In addition to standard certification tests for on-board electrical components (MIL-STD-810F, MIL-STD-461E, AGARD Flight Test Instrumentation Series), in-situ sensors must also demonstrate their reliability to detect a range of expected flaw conditions over the total expected life of the structure. From prior studies highlighting strain gauge failures during service-life, sensor redundancy and/or sensor self-monitoring/maintenance programs become necessary and must be evaluated by a validation protocol [4].

In addition to testing detection capability, avoiding false calls is a key requirement for SHM systems. Walbusser and Lindgren presented data on KC-130J exceedances resulting from sensor measurements resulting from fuel quality, accelerometers, differential pressure and position [8]. All reported KC-130J exceedances require Aircraft Structural Life engineers to assess the data and provide maintenance recommendations to the Fleet Support Team due to a high false call rate. Of 422 exceedance calls between 2001 and 2009, only 19% were considered true exceedances. Today, false calls (FC) result in added engineering costs, lead to unnecessary secondary inspections and affect aircraft availability, all detrimental to the introduction and acceptance of any new damage sensing technologies. Another example was the aforementioned in-situ damage detection systems installed on USAF KC-135 in the 1980s [7]. These systems had a high false call rate and led to significant amount of unnecessary maintenance work to determine if an indication was real damage. In today's operating environment, such false call rates would be very hard to justify. Thus, a rigorous validation protocol accurately quantifying the POD and false call rate for an SHM system is thus required to truly address these issues.

GENERAL APPROACH TO CAPABILITY VALIATION

Empirical assessment of the performance of an SHM system in its operational environment is not a trivial manner and could easily be cost prohibitive because of the extensive testing required on representative structures, expensive testing equipment, and often custom made test fixtures to establish statistically significant

performance metrics. The process defined in MIL HDBK 1823A requires the acquisition of an extensive amount of empirical data (e.g. specimens, labor, etc) , which is costly for nondestructive evaluation (NDE) systems and not realistic for an in-situ damage detection system [6]. Model-Assisted Probability of Detection (MAPOD) methods mitigate the cost of POD studies and facilitate the characterization and insertion of NDE systems. MAPOD leverages both computer models and transfer functions to enable the determination of the sensitivity of damage detection systems and the effect of changing the sensitivity threshold on the number of false calls that occur as a result of implementing an NDE system, while minimizing the need for empirical data.

A Model-Assisted Probabilistic Reliability Assessment (MAPRA) methodology for in-situ damage detection is presented, inspired by the MAPOD approach, which consists of statistical metrics of reliability for in-situ systems for damage detection, localization, and sizing, and a protocol designed for using empirical data, models, and simulations to characterize these nondestructive methods by statistical metrics, including uncertainty analysis [9-10]. This model-based methodology minimizes the number of samples that must be prepared with representative damage to obtain the data required to achieve statistically meaningful characterization results. In addition, these methods enable assessment of the effects of changing sensitivity thresholds or boundary conditions upon the number of false calls of in-situ damage detection systems and/or on the error of damage localization and sizing systems.

REQUIREMENTS FOR VALIDATION OF SHM TECHNOLOGIES

The framework for in-situ damage detection validation relies on the foundation of existing certification protocols and analysis procedures for 1) Environmental Testing of Airborne Equipment [11], 2) Materials / Structure Certification, 3) NDE (POD) Validation Procedures, 4) Failure Mode Effects and Criticality Analysis (FMECA), and 5) Cost Benefits Analysis [12]. Existing procedures for environmental testing of airborne equipment ensure flight certification for the sensor and data acquisition hardware. In addition, the certification of new materials and aircraft structure designs provides a framework for a multi-scale certification approach for in-situ damage detection applications on systems of increasing complexity. However, the critical component of reliability demonstrations is ensuring that probability of detection, false call rates, and probabilities of damage localization and sizing errors are well understood and within acceptable ranges.

It is important to recall that the capability of current periodic inspection methods to detect damage is validated by the POD process presented in MIL-HDBK-1823A [6]. The current ASIP methods to calculate risk include the incorporation of the entire POD curve (i.e. life of the component or system). Thus, point estimates of capability, such as Receiver-Operator Curves (ROC), which define sensitivity for only one damage size, are not suitable for the current ASIP approach. In addition, a statistical assumption in ROC approaches is that there is a large population of the target of interest, which is typically not the case for the damage located in USAF structures. In addition, when such a scenario is encountered, the typical approach would be to replace such a structure or change of design. Therefore, the approach leverages the use of MIL-HDBK-1823A to enable integration into current USAF practices.

The capability validation protocol includes four critical components: (1) a procedure to identify the critical factors impacting system performance; (2) a multistage or hierarchical approach to system validation (a) following the materials certification process and (b) utilizing electronics durability testing; (3) a model-assisted evaluation process to address the wide range of expected damage conditions that cannot be experimentally tested; and (4) probability of detection (POD), probability of false call (POFC) and probability of random missed call (POMC) evaluations with confidence bounds estimation and uncertainty analysis for in-situ damage detection systems, and evaluation of appropriate probabilistic metrics to characterize the quality of damage localization and damage characterization for systems that include such capabilities.

Before a reliability assessment test plan can be designed, a fundamental understanding of all the pertinent characteristics of an in-situ damage assessment system and application must be considered. These include type of damage sensing (e.g. active direct sensing, passive direct sensing, or indirect sensing via loads, or environmental data to be used by life prediction models), coverage and sensor location (e.g. local or global), time of data acquisition, (e.g. continuous or periodic), location of the DAQ hardware, type of sensing method (e.g. ultrasound or eddy current), damage type or failure conditions to detect, criticality of the damage state (e.g. safety of flight), likelihood of worst case occurrence, credit or value associated with the application, data classification approach, system maturity or technology readiness level, secondary inspection and maintenance actions, process controls, system failure modes effects analysis, system maintenance, and necessary accuracy of damage localization and characterization estimates. The validation and certification testing procedure is initiated by posing and answering questions regarding these characteristics, which guides the validation and certification tests. These answers will determine what factors must be evaluated to validate the SHM system indications

Critical Factor Assessment

Once the outline of an experimental test plan is in place, the next step is to evaluate the important factors concerning in-situ damage detection performance and reliability. Similar procedures have been developed for guiding the determination of the critical factors for NDE techniques [13]. First, the potential contributing factors must be evaluated for their role in the in-situ system performance and reliability. Such factors categories include: part geometry and material properties, loading and environmental conditions, system hardware, and the flaw characteristics. Expert opinion, prior work, baseline experimental studies and simulation can be used to make the case for each factor whether they should be considered in the full reliability study. An evaluation of each factor is needed to prescribe the type of approach that can be used to determine how each factor impacts the system performance and reliability. Determining whether either empirical and/or simulated studies are appropriate for the reliability evaluation is needed before a final validation procedure can be defined. Care must also be taken in determining the significance and role of the interaction between select factors (covariance) in the reliability study. Unique requirements for SHM systems in terms of inspection

complexity and durability indicate that great care must be taken during this initial step. Regular depot maintenance actions such as grindout of corrosion, replacement of select panels, and application of patches, will alter the dynamic characteristics of a structure, and the corresponding changes in sensor signals must be differentiated from the detection and characterization of critical defects. Other stochastic variations in the structure, due to manufacture, maintenance, repair, modification and usage, must also be addressed to determine how these random changes affect the output of the in-situ damage assessment system. If a factor is found even remotely significant on the system response, it must be considered in the study.

Multistage Testing Approach

The key characteristic of the protocol is a multistage approach to SHM system validation following the materials certification process. A similar approach has recently been proposed by the Aerospace Industry Steering Committee on Structural Health Monitoring [14]. The multistage validation approach proposes incrementally testing systems with structures of increasing complexity [15]. They include (a) coupon testing, (b) sub-component testing, (c) life testing (full-scale fatigue testing if feasible), (d) on-structure demonstration, and (e) final system verification. Each step would address a unique set of factors: (a) local damage condition with loading spectrum, (b) local sub-structure condition, (c) environmental conditions and sensor degradation, (d) global structure condition with actual system, and (e) in-service condition calibration and verification. Care must be taken in performing each study and justifying the assumptions in the multistage evaluation. In particular, false calls may not be present in the coupon testing, but will be much more problematic for each level of increasing complexity. Therefore, to address all the additional degrees of complexity, the amount of testing at the full scale level is anticipated to be much more demanding and numerous than the limited testing needed at a coupon level.

Statistical Metrics for Detection, Localization, and Characterization

As stated previously, validated capability for detection is determined via a POD-like assessment to provide the POD curve required as the input to the ASIP risk determination. However, probability of detection and probability of false calls are not sufficient descriptors of reliability for methods that provide localization and sizing of damage; several enhancements must be developed before a model assisted approach can be applied for characterizing the reliability of systems that do more than damage detection. To extend MAPRA to address damage localization and quantification cases, accurate assessments of the uncertainty bounds on the predicted values for damage position, size, and morphology are needed.

At the heart of in-situ damage localization and quantification is a fundamental parameter estimation problem. Interval estimation provides a more realistic perspective on the measures of localization and quantification. Confidence intervals must be estimated to understand the range of values that the measure represents. This is not straightforward, in particular for estimation of correlated random variables. The damage localization problem may be restated in terms of ‘probability of localization’ and damage quantification (characterization) problem can be defined on each of the elements of a characterization vector.

Model-assisted Evaluation Process

A critical part of the protocol is the option for a *model-assisted evaluation*. Global in-situ damage detection systems are particularly challenging to validate since they are designed to be sensitive to damage conditions over an entire component. Following the NDE POD experimental protocol [6], all locations of interest would require a statistically significant number of samples for validation. An in-situ POD study must test (1) that the method can reliably detect critical damage for all expected locations, and (2) that the method is subject to very low false call rate for the expected variations in operational conditions. To address both the large number of factors and the sample requirements in evaluating in-situ reliability, a model-assisted methodology becomes necessary.

A model-assisted strategy for the design and execution of POD studies for NDE has been developed and demonstrated to help mitigate validation costs and to improve POD evaluation quality by addressing a wider array of inspection variables. At one level, empirically developed models can be used to transform POD results from one set of conditions to another. For example, a transfer function model-based POD evaluation has been successfully demonstrated for detecting cracks in engine components [16] and aircraft structures [17]. In addition, progress has also been made on a full model-assisted POD methodology using numerical simulation. Recently, demonstrations have been made incorporating computer simulation for the inspection of cracks around fastener sites in a two layer aircraft wing-type structure inspection performed with an eddy current technique [18]. A unified protocol for performing model-assisted POD assessment has been defined and recent application results were highlighted [5].

To address the challenge of limited data, a model-based assessment must be extended, beyond a simple deterministic representation of a nondestructive measurement, to incorporate the variations of the most significant input factors and include uncertainty propagation in the evaluation. Developing a validated stochastic model that includes all significant sources of variation on the measurement response is key to sample and experimental test reduction. Simultaneously, less costly studies could be performed to quantify precisely the probability density functions (with confidence bounds) for the key controlling factors.

CLOSING REMARKS

The drivers and requirements for validation of in-situ damage assessment systems have been described and a framework for in-situ methods was defined. The next steps will be to verify the damage detection validation protocol using an in-situ system of interest, and to extend the methodology to validate in-situ systems that localize and characterize damage, based on the probabilistic approach outlined here.

REFERENCES

1. Fitzwater, L. M., Davis, C. L., Torng, T. Y., “Integration of Non-Deterministic Analysis and In-situ Monitoring for Structural Integrity,” *2011 Aircraft Airworthiness & Sustainment Conference*, San Diego, CA, (April 2011).

2. Ball, D. L., Lougheed, J. C., Ingram, J. E., and Burke, T. P., "Integration of Non-Deterministic Analysis and In-situ Monitoring for Structural Integrity," *2011 Aircraft Airworthiness & Sustainment Conference,* San Diego, CA, (April 2011).
3. Babish, C. A., "Requirements Associated with Transition of ISHM in USAF Aircraft," *Integrated Systems Health Management (ISHM) Conference Proceedings,* Covington, KY, (August 2009).
4. Aldrin, J. C., Medina, E. A., Lindgren, E. A., Buynak, C. F., Steffes, G., Derriso, M., "Model-assisted Probabilistic Reliability Assessment for Structural Health Monitoring Systems," *Review of Progress in QNDE*, Vol. 29, AIP, pp. 1965-1972, (2010).
5. R. B. Thompson, "A unified approach to the model-assisted determination of probability of detection", *Materials Evaluation*, Vol. 66, pp. 667-673, (2008).
6. U.S. Department of Defense, Handbook, Nondestructive Evaluation System Reliability Assessment, MIL-HDBK-1823A, 7 April 2009.
7. Bakse, C., *Airlift Tanker: History of U.S. Airlift and Tanker Forces*, Turner Publishing, (1996), p63.
8. Lindgren. E. and Walbusser. R., "Experience/Lessons Learned using Flight Parameter Sensors on US Department of Navy C-130 Aircraft", Aircraft Airworthiness and Sustainment, (Austin, TX, 2010).
9. Lindgren, E. A., Buynak, C. F., Aldrin, J. C., Median, E. A., Derriso, M. M., "Model-assisted Methods for Validation of Structural Health Monitoring Systems" *Proceedings of the 7th International Workshop on Structural Health Monitoring,* Ed. F.-K. Chang, Stanford, CA (September 9-11, 2009).
10. Aldrin, J. C., Medina, E. A., Lindgren, E. A., Buynak, C. F., Steffes, G., Derriso, M., "Model- assisted Probabilistic Reliability Assessment for Structural Health Monitoring Systems," *Review of Progress in QNDE*, Vol. 29, AIP, pp. 1965-1972, (2010).
11. Kessler, S.S., "Certifying a Structural Health Monitoring System: Characterizing Durability, Reliability and Longevity," *Proceedings of the 1st International Forum on Integrated Systems Health Engineering and Management in Aerospace*, Napa, CA, 7-10 November 2005.
12. Medina, E.A. and Aldrin, J.C. "Value Assessment Approaches for Structural Life Management through SHM, in *Encyclopedia of Structural Health Monitoring*, Boller, C., Chang, F., and Fujino, Y. (eds.) John Wiley & Sonds Ltd., Chichester, UK, pp. 1673-1690, (2009).
13. Berens, A., Hovey, P, and Skinn, D., "Risk Analysis for Aging Aircraft Fleets, Volume 1 – Analysis," UDRI, Final Report for Period Sep. 1987 – Jan. 1991, Contract F33615-87-C-3215, to Flight Dynamics Directorate, Wright Laboratory, Wright-Patterson AFB.
14. Chang, F. K., "Aerospace Industry Steering Committee on Structural Health Monitoring," AFOSR SHM Workshop, (Covington, KY, August 11, 2008).
15. Hasselman, T. K., Wathugala, G. W., "A Hierachical Approach for Model Validation and Uncertainty Quantification," Fifth World Congress on Computational Mechanics, Vienna Austria, July 7-12, (2002).
16. Smith, K. D., Thompson, R. B. and Brasche, L., "Model-Based POD: Successes and Opportunities," 1st Meeting of the MAPOD Working Group, Albuquerque, New Mexico, (September 23-24, 2004), web site: http://www.cnde.iastate.edu/MAPOD/.
17. Harding, C., G. Hugo and S. Bowles, "Model-assisted Probability of Detection Validation of Automated Ultrasonic Scanning for Crack Detection at Fastener Holes," Proceedings of the 10th Joint FAA/DoD/NASA Conference on Aging Aircraft, Palm Springs, CA, (April 16-19, 2007).
18. Aldrin, J.C., Knopp, J. S., Lindgren, E. A., Jata, K. V., "Model-assisted Probability of Detection (MAPOD) Evaluation for Eddy Current Inspection of Fastener Sites", *Review of Progress in QNDE*, (to be published, 2009).